Here I define the Bayesian model I want to use to test the performance of the different priors. The features I want to include are the folowing: continuous outcome, any number/measurement scale.

Let's start with the **model**

$$y_{ij} = \boldsymbol{x}_{ij}^T\boldsymbol{\theta} + \boldsymbol{z}_{ij}^T\boldsymbol{b}_i + \epsilon_{ij} \tag{1}$$

$\boldsymbol{b}_i \sim N(\boldsymbol{0}, \boldsymbol{\Psi})$
$\epsilon_{ij} \sim N(0, \sigma^2)$

(Vectors are in bold, amtrix are capital greek letters).
I'm going to define the following **priors**:

$$p(\boldsymbol{\theta}) \propto 1 \tag{2}$$

$$p(\sigma^2) \propto \sigma^{-2} \tag{3}$$

$$p(\boldsymbol{\Psi}) \propto [...] \tag{4}$$

The derivation of the conditional posterior follows.

Full conditional for $\boldsymbol{\theta}$ (fixed effects)

Let's start with

$$p(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{b}_i, \boldsymbol{\Psi}, \sigma^2) = p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{b}_i, \boldsymbol{\Psi}, \sigma^2)p(\boldsymbol{\theta}) \tag{5}$$

where

$p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\Psi}, \sigma^2) = \prod_{i=1}^n \prod_{j=1}^J p(y_{ij}|\boldsymbol{\theta}^T\boldsymbol{x}_{ij} + \boldsymbol{b}_i^T\boldsymbol{z}_{ij}, \boldsymbol{\Psi}, \sigma^2) \propto exp(-\frac{1}{2\sigma^2}SSR)$

and

$SSR = \sum_{i=1}^n [\sum_{j=1}^J (y_{ij} - \boldsymbol{\theta}^T\boldsymbol{x}_{ij} - \boldsymbol{b}_i^T\boldsymbol{z}_{ij})^2]$
where can rewrite y_{ij} as $\tilde{y}_{ij}$, with $\tilde{y}_{ij} = y_{ij} - \boldsymbol{b}_i^T\boldsymbol{z}_{ij}$. This would turn SSR in:

$SSR = \sum_{i=1}^n [\sum_{j=1}^J (\tilde{y}_{ij} - \boldsymbol{\theta}^T\boldsymbol{x}_{ij})^2] =$
$= (\tilde{\boldsymbol{y}} - \boldsymbol{X\theta})^T(\tilde{\boldsymbol{y}} - \boldsymbol{X\theta})$
$= \tilde{\boldsymbol{y}}^T\tilde{\boldsymbol{y}} - 2\boldsymbol{\theta}^T\boldsymbol{X}\tilde{\boldsymbol{y}} + \boldsymbol{\theta}^T\boldsymbol{X}^T\boldsymbol{X\theta}$

Hence,

$$p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\Psi}, \sigma^2) \propto exp(-\frac{1}{2\sigma^2}[-2\boldsymbol{\theta}^T\boldsymbol{X}\tilde{\boldsymbol{y}} + \boldsymbol{\theta}^T\boldsymbol{X}^T\boldsymbol{X\theta}]) \tag{6}$$

Combining this with the prior we obtain:

$$p(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\Psi}, \sigma^2) \propto exp(-\frac{1}{2}\boldsymbol{\theta}^T \boldsymbol{X}^T \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\theta}^T \boldsymbol{X}\tilde{\boldsymbol{y}}) \tag{7}$$

$$p(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\Psi}, \sigma^2) \sim multivariate - N(\frac{(\boldsymbol{X}^T \boldsymbol{X})^{-1}\boldsymbol{X}\tilde{\boldsymbol{y}}}{\sigma^2}, \frac{(\boldsymbol{X}^T \boldsymbol{X})^{-1}}{\sigma^2}) \tag{8}$$

Full conditional for $\boldsymbol{b}_i$ (random effects)

To derive this one we can start from:

$$p(\boldsymbol{b}_i|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\theta}, \boldsymbol{\Psi}, \sigma^2) = p(\boldsymbol{y_i}|\boldsymbol{\theta}, \boldsymbol{b}_i, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\Psi}, \sigma^2)p(\boldsymbol{b}_i) \tag{9}$$

We know that

$p(\boldsymbol{y}_i|.) = \prod_{j=1}^J p(y_{ij}|\boldsymbol{\theta}^T \boldsymbol{x}_{ij} + \boldsymbol{b}_i^T \boldsymbol{z}_{ij}, \sigma^2) \propto exp(-\frac{1}{2\sigma^2}SSR_i)$

with

$SSR = \sum_{j=1}^J (y_{ij} - \boldsymbol{\theta}^T \boldsymbol{x}_{ij} - \boldsymbol{b}_i^T \boldsymbol{z}_{ij})^2$

and we can rewrite y_{ij} as $\tilde{y}_{ij} = y_{ij} - \boldsymbol{\theta}^T \boldsymbol{x}_{ij}$, which would make SSR be

$SSR = \sum_{j=1}^J (\tilde{y}_j - \boldsymbol{\theta}^T \boldsymbol{x}_j)^2 =$
$= (\tilde{\boldsymbol{y}} - \boldsymbol{b}_i^T \boldsymbol{Z}_i)^T (\tilde{\boldsymbol{y}} - \boldsymbol{b}_i^T \boldsymbol{Z}_i)$
$= \tilde{\boldsymbol{y}}^T \tilde{\boldsymbol{y}} - 2\boldsymbol{b}_i^T \boldsymbol{Z}_j \tilde{\boldsymbol{y}} + \boldsymbol{b}_i^T \boldsymbol{Z}_i^T \boldsymbol{Z}_j \boldsymbol{b}_i$

Hence,

$$p(\boldsymbol{y}_i|.) \propto exp(-\frac{1}{2\sigma^2}[-2\boldsymbol{b}_i^T \boldsymbol{Z}_j \tilde{\boldsymbol{y}} + \boldsymbol{b}_i^T \boldsymbol{Z}_i^T \boldsymbol{Z}_j \boldsymbol{b}_i]) \tag{10}$$

We also know that in this case, the "prior" is

$$p(\boldsymbol{b}_i) \propto N(\boldsymbol{0}, \boldsymbol{\Psi}) \propto exp(-\frac{1}{2}[-2\boldsymbol{b}_i^T \boldsymbol{\Psi}^{-1}\boldsymbol{0} + \boldsymbol{b}_i^T \boldsymbol{\Psi}^{-1}\boldsymbol{b}_i]) \tag{11}$$

In conclusion, combining the sampling model and the prior, we get:

$$p(\boldsymbol{b}_i|.) \propto exp(-\frac{1}{2\sigma^2}[-2\boldsymbol{b}_i^T \boldsymbol{Z}_j \tilde{\boldsymbol{y}} + \boldsymbol{b}_i^T \boldsymbol{Z}_i^T \boldsymbol{Z}_j \boldsymbol{b}_i] - \frac{1}{2\sigma^2}[-2\boldsymbol{b}_i^T \boldsymbol{Z}_j \tilde{\boldsymbol{y}} + \boldsymbol{b}_i^T \boldsymbol{Z}_i^T \boldsymbol{Z}_j \boldsymbol{b}_i]) \tag{12}$$

$$p(\boldsymbol{b}_i|.) \propto multiv-N((\Psi^{-1} + \frac{\boldsymbol{Z}_i^T \boldsymbol{Z}_i}{\sigma^2})^{-1}(\boldsymbol{\Psi}^{-1}\boldsymbol{0} + \frac{\boldsymbol{Z}_i^T \tilde{y}_i}{\sigma^2}), (\Psi^{-1} + \frac{\boldsymbol{Z}_i^T \boldsymbol{Z}_i}{\sigma^2})^{-1}) \tag{13}$$

Full conditional for σ^2 (error variance) The full conditional posterior can be expressed as:

$$p(\sigma^2|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\theta}, \boldsymbol{b}_i, \boldsymbol{\Psi}) = p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{b}_i, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\Psi}, \sigma^2)p(\sigma^2) \tag{14}$$

The sampling model is the same we saw for the full conditional distribution of $\boldsymbol{\theta}$:

$p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\Psi}, \sigma^2) = \prod_{i=1}^{n} \prod_{j=1}^{J} p(y_{ij}|\boldsymbol{\theta}^T \boldsymbol{x}_{ij} + \boldsymbol{b}_i^T \boldsymbol{z}_{ij}, \boldsymbol{\Psi}, \sigma^2) =$
$= \prod_{i=1}^{n} \prod_{j=1}^{J} (2\pi\sigma^{-2})^{-\frac{1}{2}} exp(-\frac{(y_{ij} - \boldsymbol{\theta}^T \boldsymbol{x}_{ij} - \boldsymbol{b}_i^T \boldsymbol{z}_{ij})^2}{2\sigma^2})$

However, we are now interested in σ^2, hence

$p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\Psi}, \sigma^2) \propto (\sigma^2)^{-\frac{N}{2}} exp(-\frac{\sum_{i=1}^{n} \sum_{j=1}^{J} (y_{ij} - \boldsymbol{\theta}^T \boldsymbol{x}_{ij} - \boldsymbol{b}_i^T \boldsymbol{z}_{ij})^2}{2\sigma^2}) = (\sigma^2)^{-\frac{N}{2}} exp(-\frac{1}{2\sigma^2} SSR)$

where $N = \sum_{i}^{n} nj_i$ is the entire sample size (all observations within all clusters).

The prior for σ is given above, and therefore we can write the full conditional posterior as:

$$p(\sigma^2|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\theta}, \boldsymbol{b}_i, \boldsymbol{\Psi}) \propto (\sigma^2)^{-\frac{N}{2}-1} exp(-\frac{1}{2\sigma^2} SSR) \tag{15}$$

$$p(\sigma^2|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\theta}, \boldsymbol{b}_i, \boldsymbol{\Psi}) \sim IG(\frac{N}{2}, \frac{SSR}{2}) \tag{16}$$

Full conditional for $\boldsymbol{\Psi}$ (random effects variance covariance matrix)

See Mulder Pericchi, 2018

Notation Conventions

- n number of clusters; i specific cluster
- J number of observations within cluster; j specific observation
- N total number of observations